

Curso Inteligência Artificial: do Zero ao Infinito

Introdução à Regressão Linear

Universidade Federal de Mato Grosso

Agenda

1 Introdução

2 Regressão Linear

O que aprenderei neste Curso?

Módulo Intermediário

- Introdução às Redes Neurais proporcionando ao aluno uma base em aprendizagem profunda e redes neurais.
- Implementação de gradiente descendente e retropropagação em Python. Treinamento de redes neurais proporcionando ao aluno o aprendizado de técnicas para melhorar o treinamento de uma rede neural, tais como: early-stopping, regularização e dropout.
- Aprender os princípios básicos das camadas de uma rede neural convolucional (CNN): convolutional, maxpooling e fully-connected.
- Apresentação dos princípios básicos sobre as principais arquiteturas CNN para detecção e segmentação de imagens.
- Duração Estimada: 5 aulas

Objetivos

Na aula de hoje, nosso objetivo é:

- Aprender os conceitos básicos sobre **Regressão linear**.
- Realizar experimentos utilizando a biblioteca *SKlearn*.
- Primeiro contato com a biblioteca *Tensorflow-Keras*.

Regressão Linear

- Regressão Linear é dos primeiros modelos estatísticos que se tem contato ao estudar *machine learning*.
- É uma equação para se estimar um alvo (variável y ou dependente), dados os valores de outras variáveis (variáveis x ou independentes).
- Um algoritmo simples, mas que usado com os parâmetros corretos é capaz de oferecer uma grande capacidade preditiva somente com a relação das suas variáveis.

Regressão Linear

- A regressão linear pode ser definida pela estatística como uma equação que busca estimar o(s) valor(es) de y , dados uma ou mais variáveis x .

$$y = a + bX$$

- Assim y é a variável dependente de x , a é o coeficiente linear e b é o coeficiente angular.

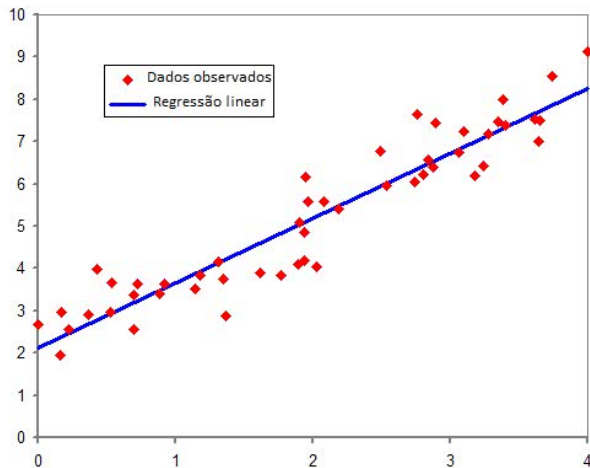
Regressão Linear

Cálculo de a e b :

$$b = \frac{\text{cov}(X, Y)}{\text{var}(X)}$$

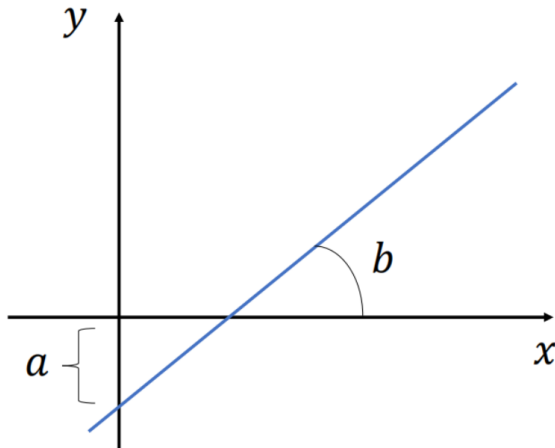
$$a = \bar{Y} - b\bar{X}$$

Regressão Linear



Regressão Linear

- A variável a que vai definir o deslocamento da reta e a variável b que vai definir a inclinação da reta.



Regressão Linear Múltipla

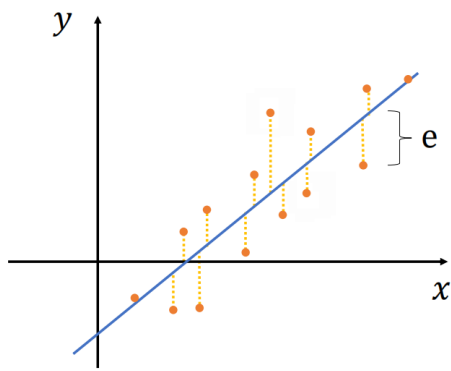
Na Regressão Linear Múltipla há mais variáveis, gerando a equação:

$$y = b_0 + b_1X_1 + b_2X_2 + \dots + b_kX_k + e$$

em que e é o erro aleatório.

Regressão Linear

- O erro é dado por e , que é a diferença entre a saída original e a saída prevista pelo modelo



- Essa diferença pode ser positiva ou negativa e isso vai influenciar na análise de erro, por isso é normal analisar o *erro médio absoluto*, o *erro quadrático médio* e o *coeficiente de determinação*.

Regressão Linear

Erro Médio Absoluto

- O erro médio absoluto (MAE) é a média da soma de todos os erros;
- A sua análise sofre uma interferência devido aos erros positivos e negativos se anularem.

Regressão Linear

Erro Quadrado Médio

- O erro quadrado médio (MSE) é a média da soma de todos os erros elevados ao quadrado;
- As diferenças elevadas ao quadrados resolve o problema de os erros positivos e negativos se anulam, sendo mais preciso que o MAE.

Regressão Linear

Coeficiente de Determinação (R^2 Score)

- O coeficiente de Determinação (R^2) varia entre 0 e 1 e expressa a quantidade da variância dos dados que é explicada pelo modelo linear.
- Por exemplo, $R^2 = 0,56$ significa que o modelo consegue prever 56% da variância total dos dados.

Referências

- Análise de Regressão Linear
- Introduction to Machine Learning Algorithms: Linear Regression
- Tutorial: Understanding Regression Error Metrics in Python

Curso Inteligência Artificial: do Zero ao Infinito

Introdução à Regressão Linear

Universidade Federal de Mato Grosso